
Supplementary information

**Metabolomic profiles predict individual
multidisease outcomes**

In the format provided by the
authors and unedited

Supplementary Note 1

TRIPOD Checklist: Prediction Model Development and Validation



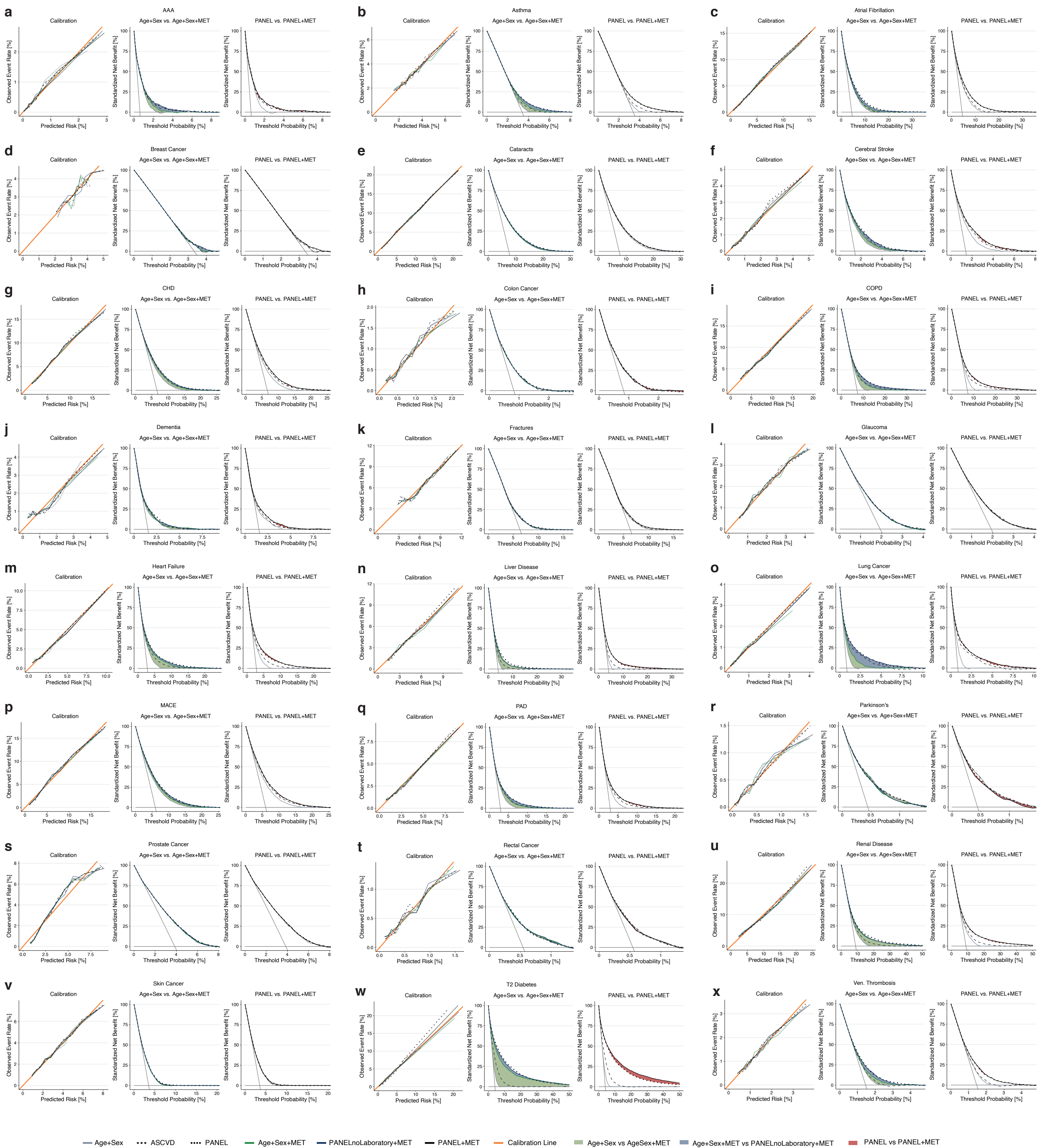
Section/Topic	Item	Checklist Item	Page
Title and abstract			
Title	1	D;V Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract	2	D;V Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	4
Introduction			
Background and objectives	3a	D;V Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	5ff
	3b	D;V Specify the objectives, including whether the study describes the development or validation of the model or both.	5ff
Methods			
Source of data	4a	D;V Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	5ff, 37ff
	4b	D;V Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	37ff
Participants	5a	D;V Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	6+7, 37ff
	5b	D;V Describe eligibility criteria for participants.	6+7, 37ff
	5c	D;V Give details of treatments received, if relevant.	37ff
Outcome	6a	D;V Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	37ff, Supplementary Table 1
	6b	D;V Report any actions to blind assessment of the outcome to be predicted.	NA
Predictors	7a	D;V Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	Supplementary Table 1,2,3
	7b	D;V Report any actions to blind assessment of predictors for the outcome and other predictors.	NA
Sample size	8	D;V Explain how the study size was arrived at.	NA
Missing data	9	D;V Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	37+38
Statistical analysis methods	10a	D Describe how predictors were handled in the analyses.	38, 39, 40
	10b	D Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	5-8, 38ff
	10c	V For validation, describe how the predictions were calculated.	41+42
	10d	D;V Specify all measures used to assess model performance and, if relevant, to compare multiple models.	9,10,11, 40+41
	10e	V Describe any model updating (e.g., recalibration) arising from the validation, if done.	41
Risk groups	11	D;V Provide details on how risk groups were created, if done.	12+13
Development vs. validation	12	V For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	41
Results			
Participants	13a	D;V Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	6,7,8, Fig1
	13b	D;V Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	6,7,8, Table 1
	13c	V For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	Suppl. Table 4
Model development	14a	D Specify the number of participants and outcome events in each analysis.	Suppl. Table 1
	14b	D If done, report the unadjusted association between each candidate predictor and outcome.	NA
Model specification	15a	D Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	39, Github
	15b	D Explain how to use the prediction model.	39, Github
Model performance	16	D;V Report performance measures (with CIs) for the prediction model.	Fig. 3,4 Extended Data Fig. 4
Model-updating	17	V If done, report the results from any model updating (i.e., model specification, model performance).	NA
Discussion			
Limitations	18	D;V Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	17ff
Interpretation	19a	V For validation, discuss the results with reference to performance in the development data, and any other validation data.	18, 17
	19b	D;V Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	17ff
Implications	20	D;V Discuss the potential clinical use of the model and implications for future research.	17ff
Other information			
Supplementary information	21	D;V Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	20
Funding	22	D;V Give the source of funding and the role of the funders for the present study.	20ff

*Items relevant only to the development of a prediction model are denoted by D, items relating solely to a validation of a prediction model are denoted by V, and items relating to both are denoted D;V. We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.

Supplementary Figure 1

Supplementary Figure 1: Model calibration and clinical utility for all endpoints

a-x) Displayed are calibration curves and clinical utility decision curves for all endpoints. In each panel, the left figure displays the calibration curve for Cox proportional hazard models, including the baseline parameter sets AgeSex, ASCVD, and PANEL, as well as their extensions with the metabolomic state (i.e. AgeSex+MET). All models are well-calibrated. The center figure displays endpoint-specific net benefit curves standardized by the endpoint prevalence, where the horizontal solid gray line indicates “treat none” and the vertical solid gray line indicates “treat all”. The standardized net benefit of AgeSex (gray line), ASCVD (dashed dark gray line) and the PANEL set (dotted black line), are compared to AgeSex+MET (green line) and additional non-laboratory predictors of the PANEL (PANELnoLaboratory, blue line). The green alpha indicates the added benefit of the metabolomic state addition over AgeSex, while the blue alpha indicates added performance over PANELnoLaboratory, respectively. The right figure displays the standardized net benefit curves comparing the performance of PANEL+MET (solid black line) against all baselines: AgeSex (gray line), ASCVD (dashed dark gray line), and the PANEL set (dotted black line).

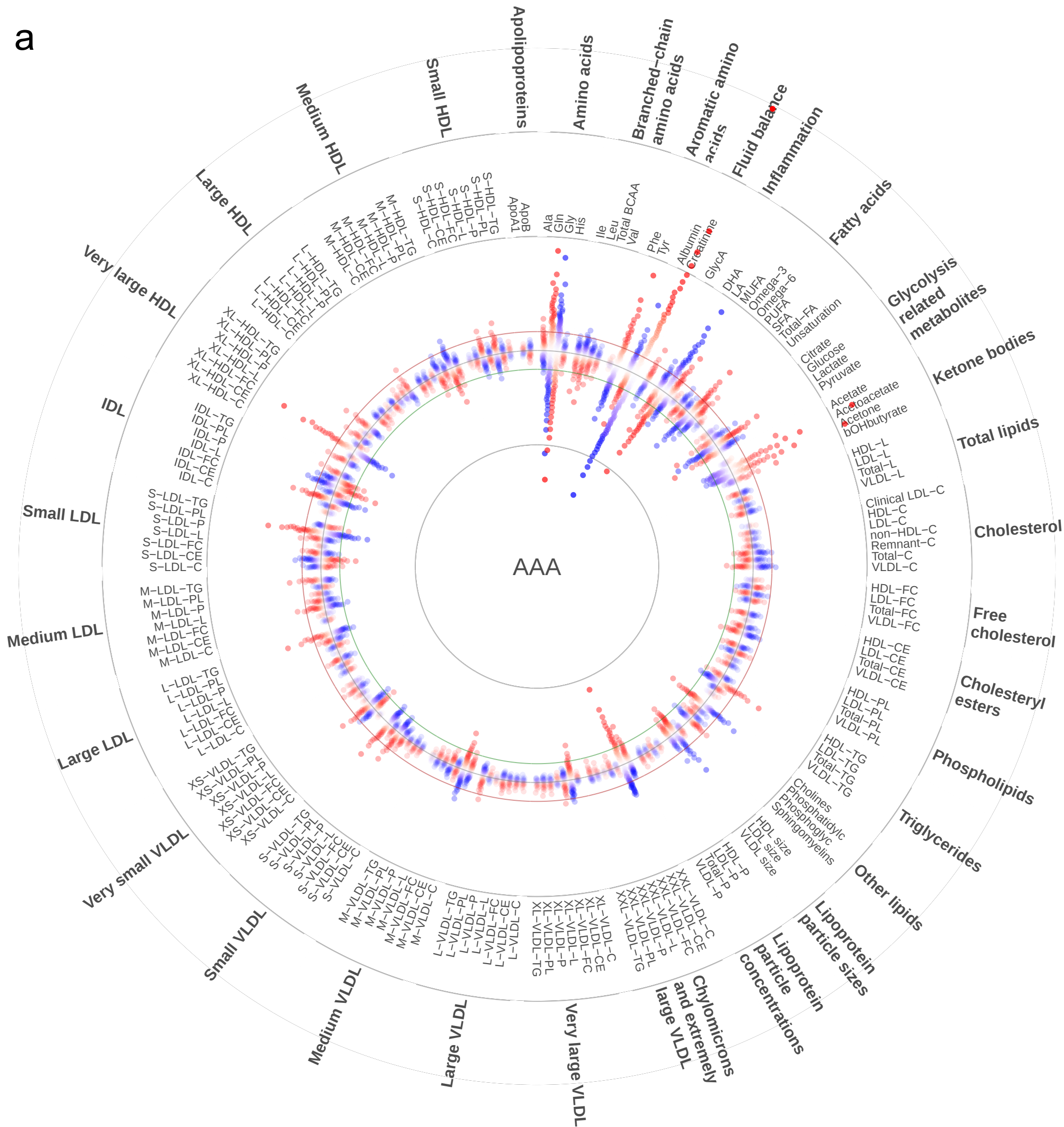


Supplementary Figure 2

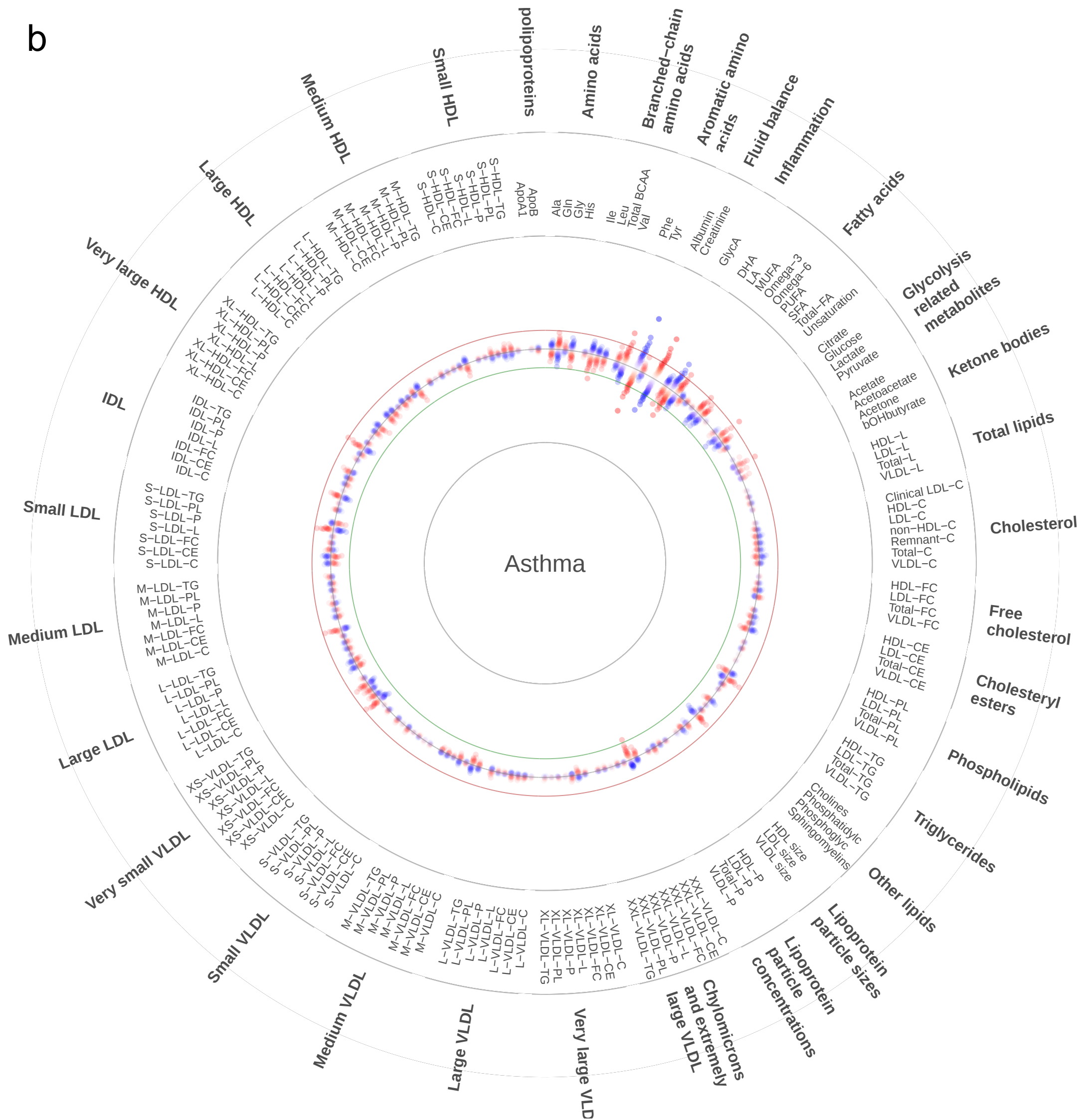
Supplementary Figure 2: Global attributions for all endpoints

a-x) Global metabolite attributions for all endpoints. Individual attributions are aggregated by percentiles. Each dot indicates one percentile. The more distant a dot from the circular baseline, the stronger the absolute attribution for the percentile. Deviations toward the center represent negative, and deviations toward the outside represent positive contributions to the metabolomic state. The color indicates the metabolite's mean normalized plasma value.

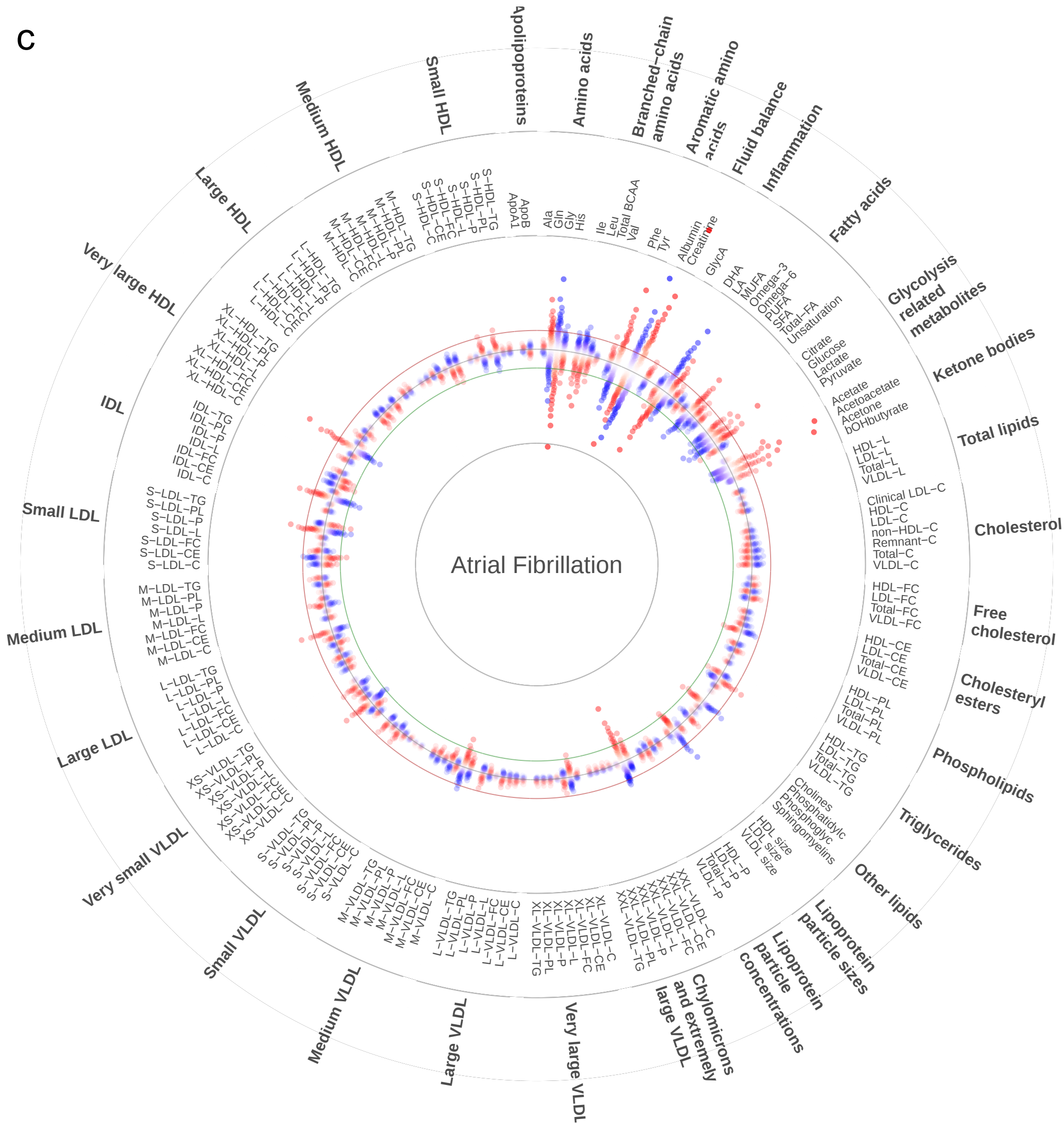
a



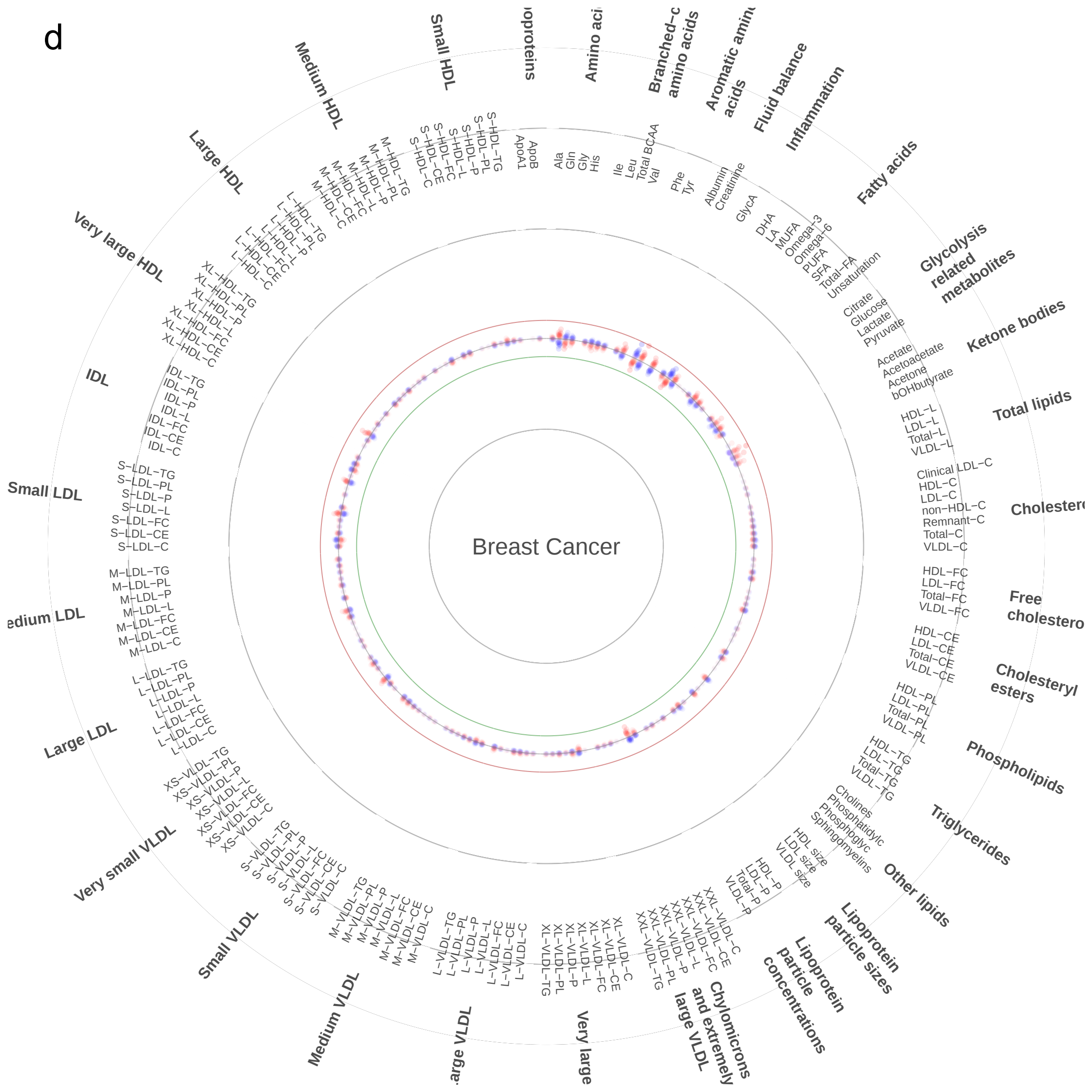
b



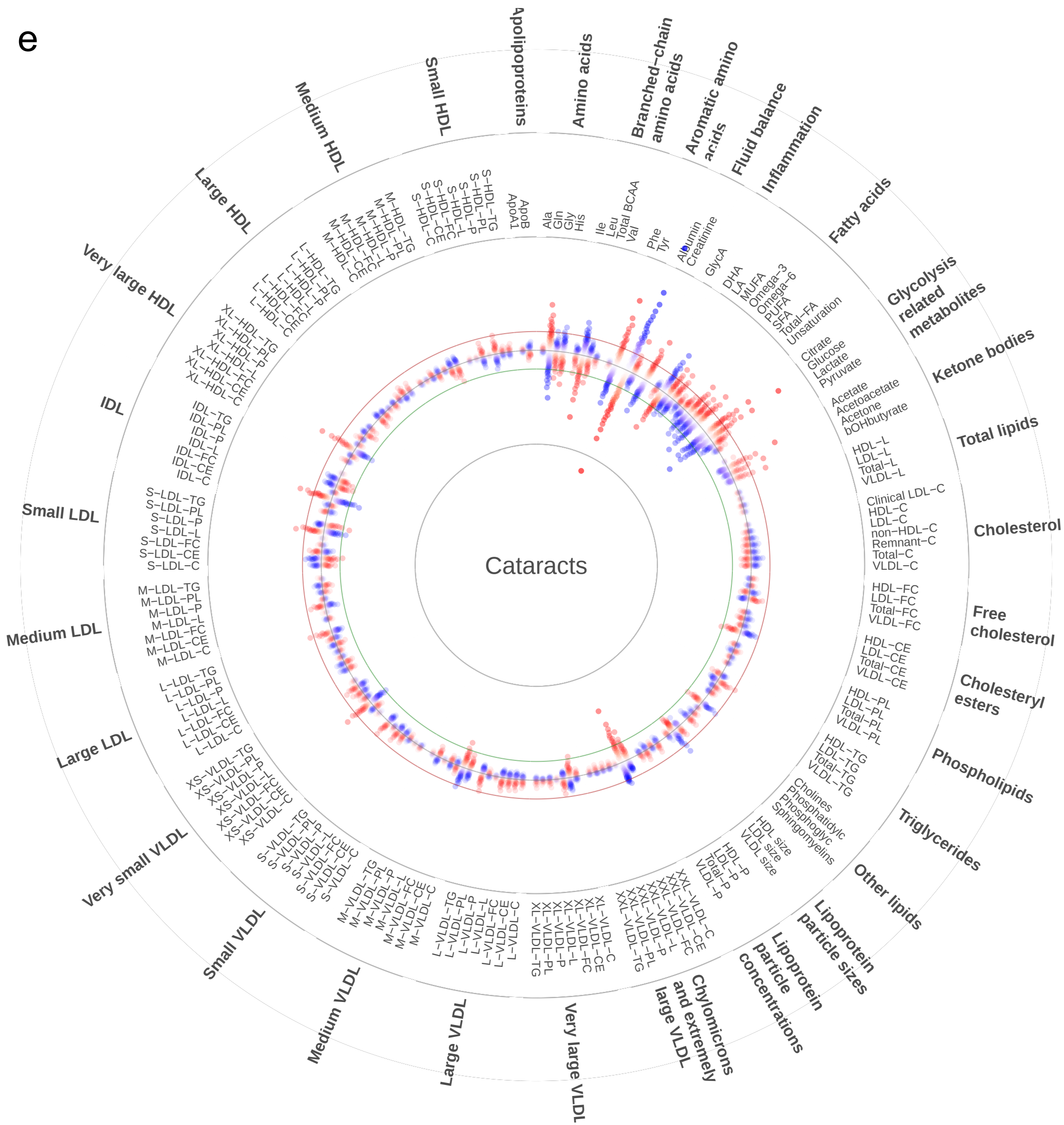
C



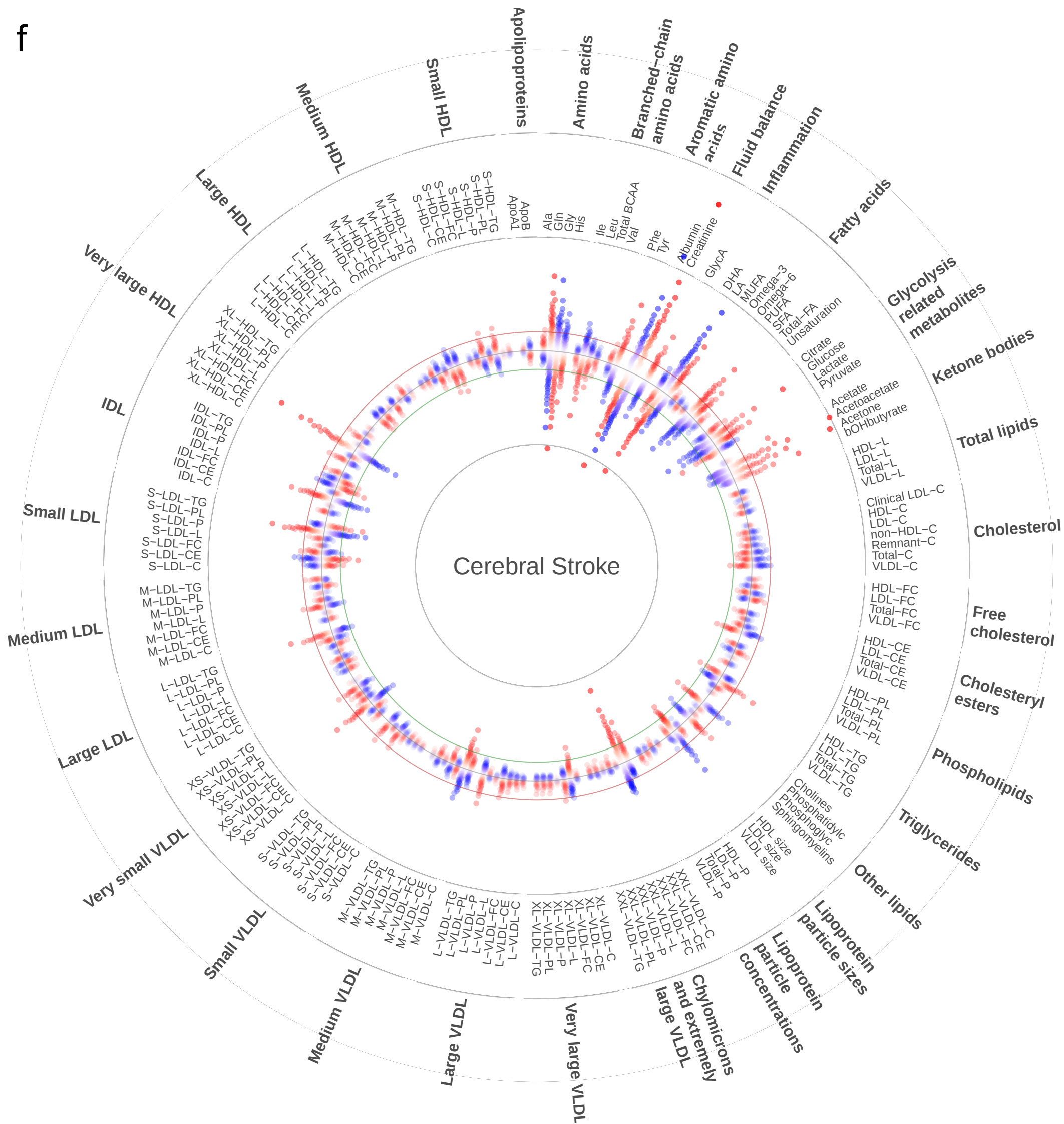
d

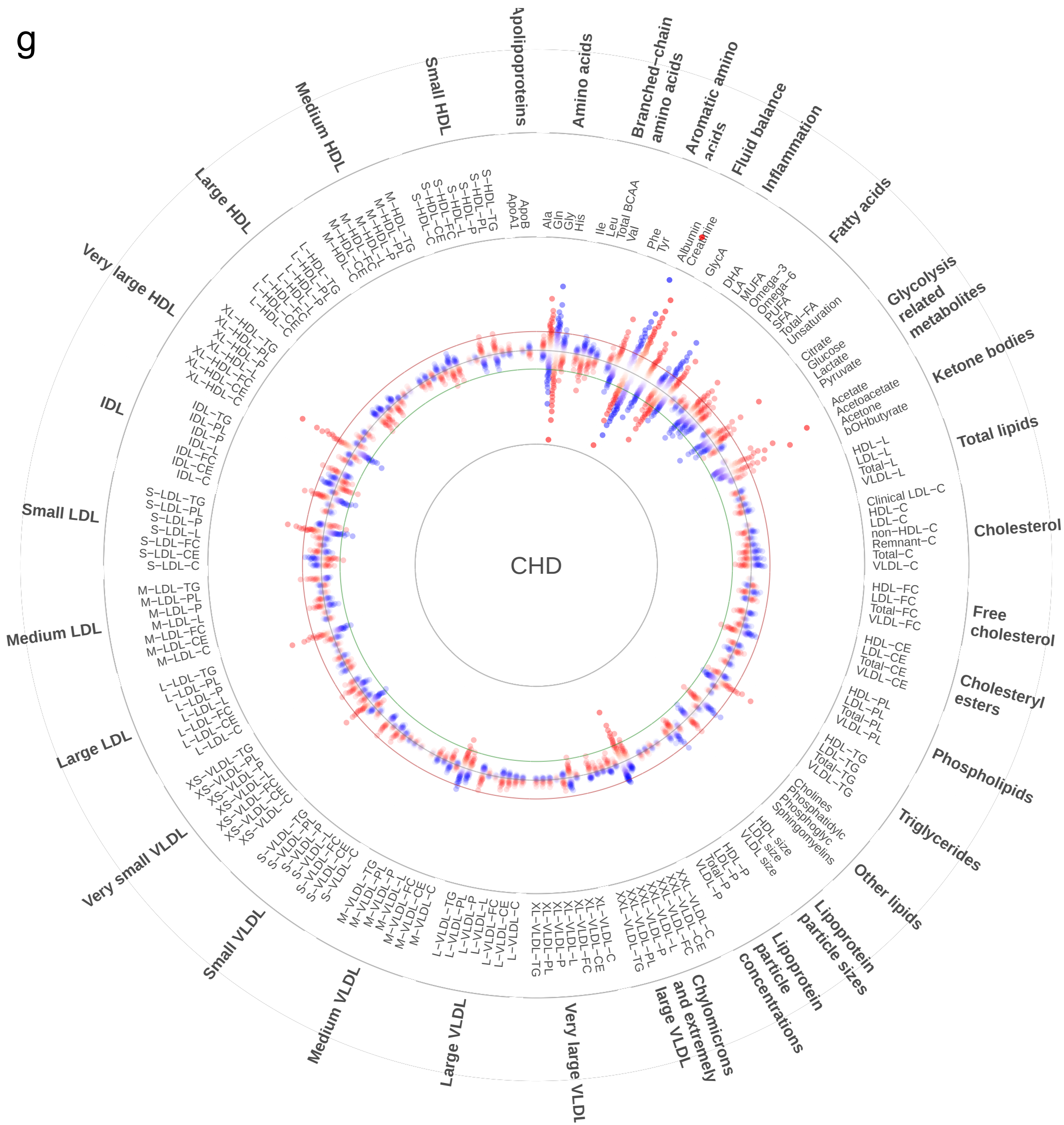


e

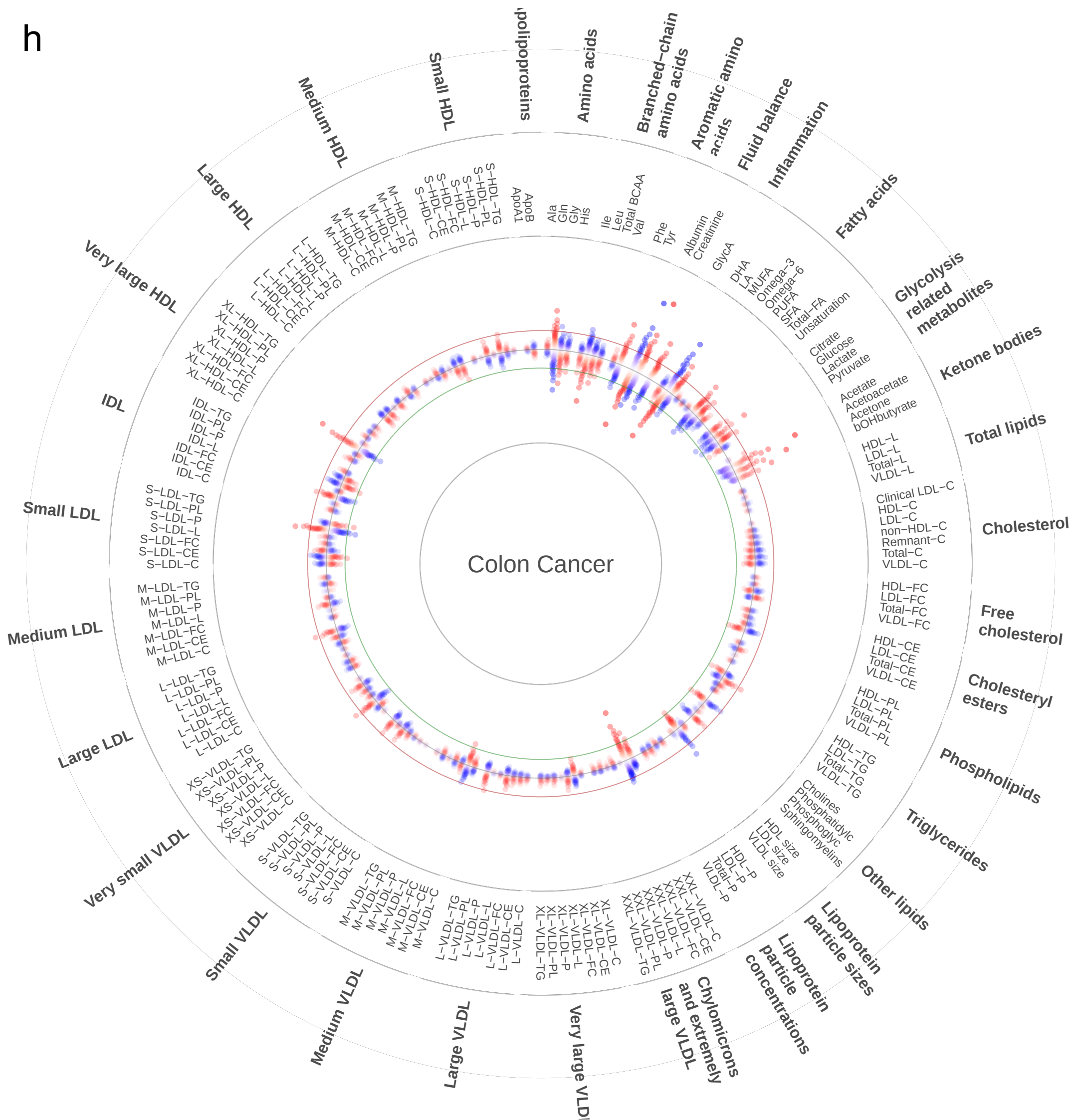


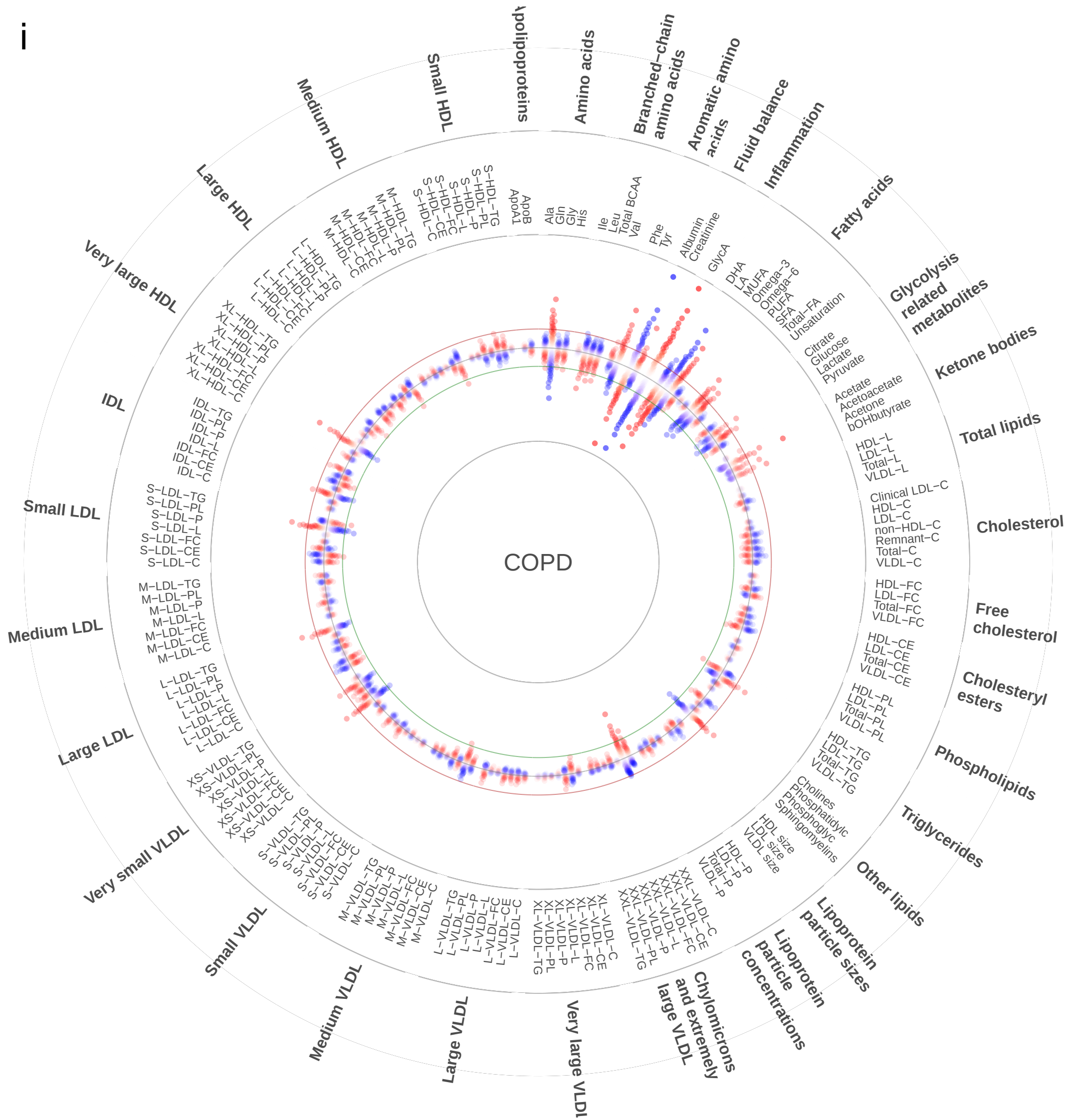
f



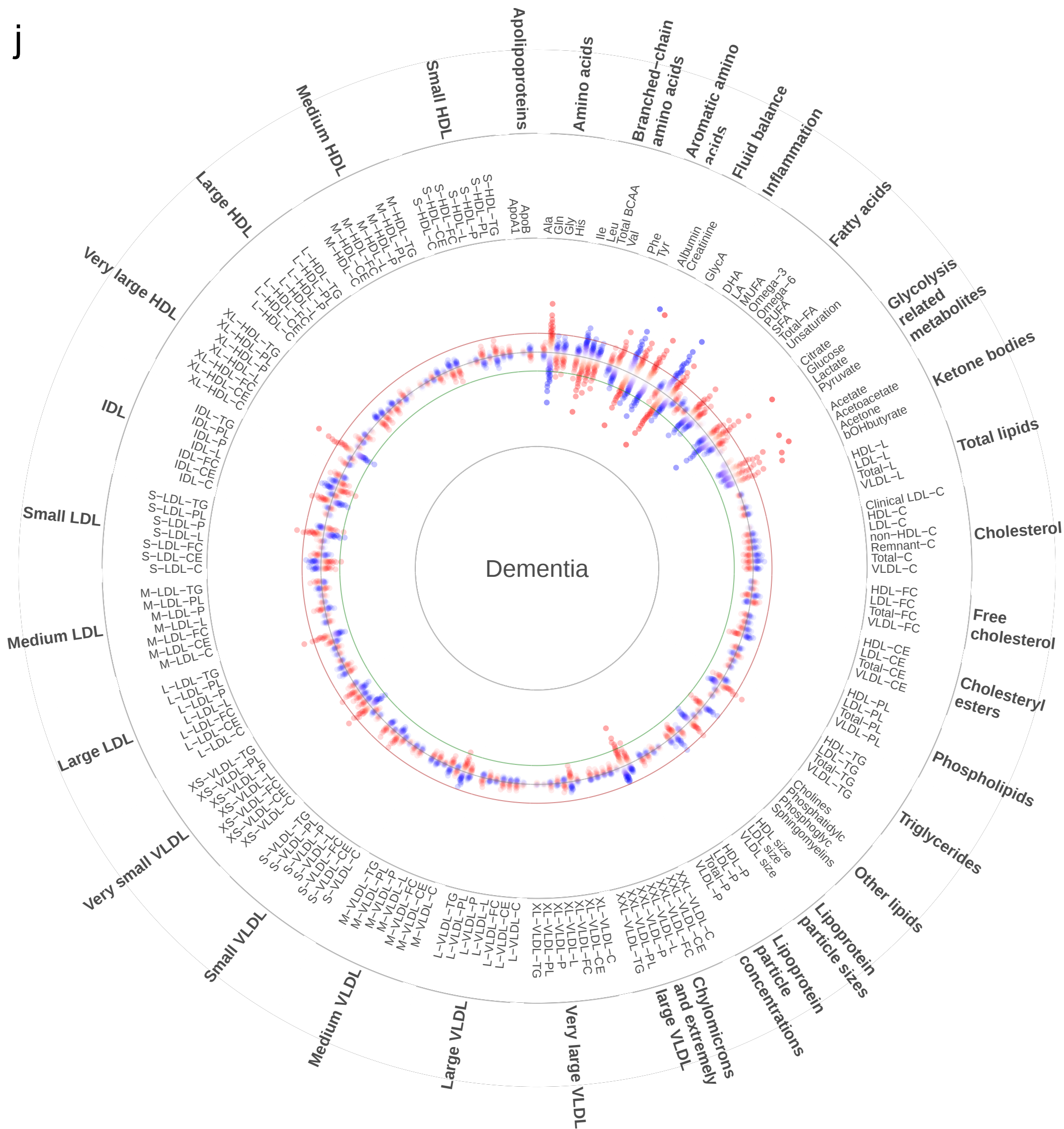


h

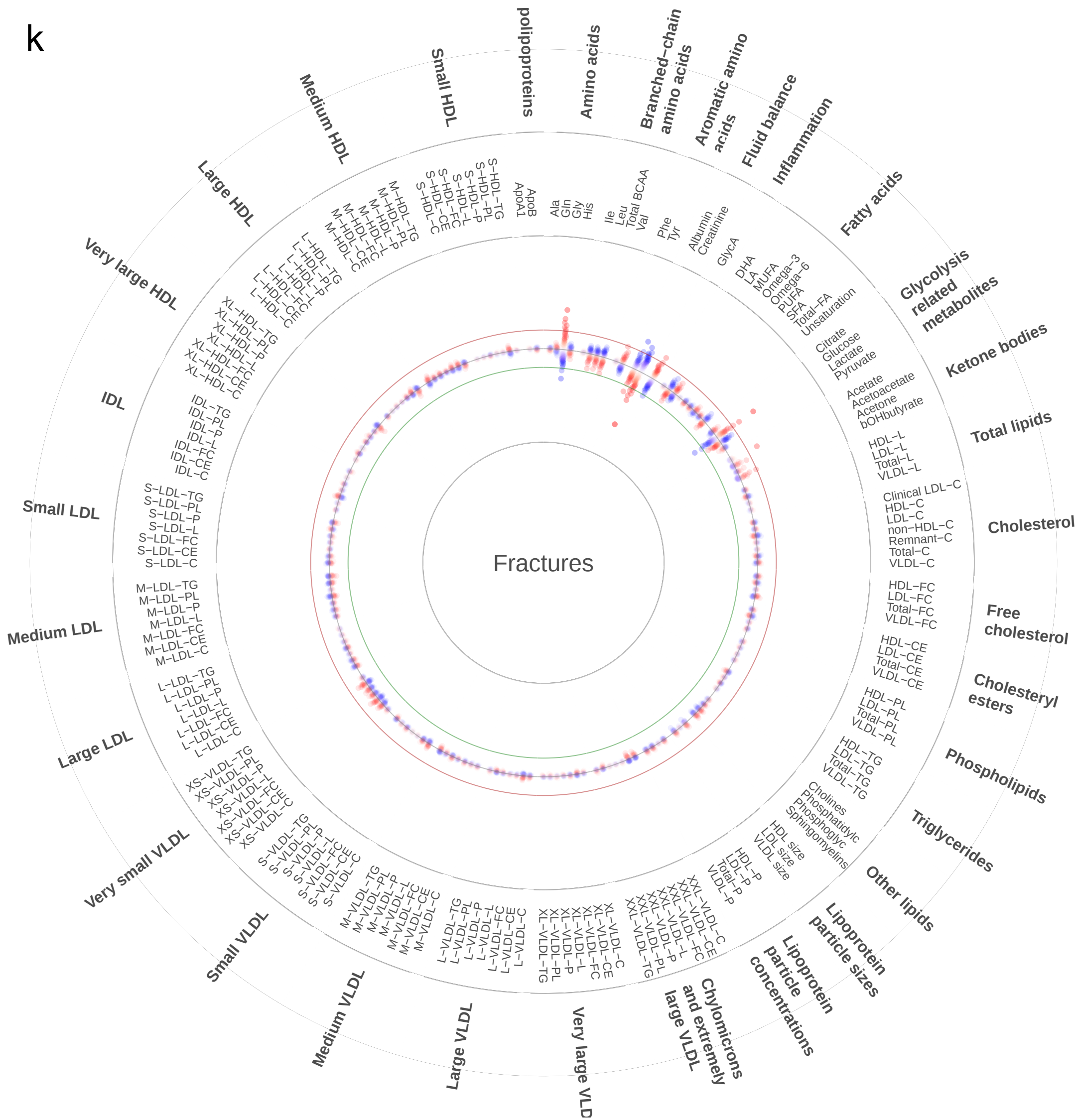


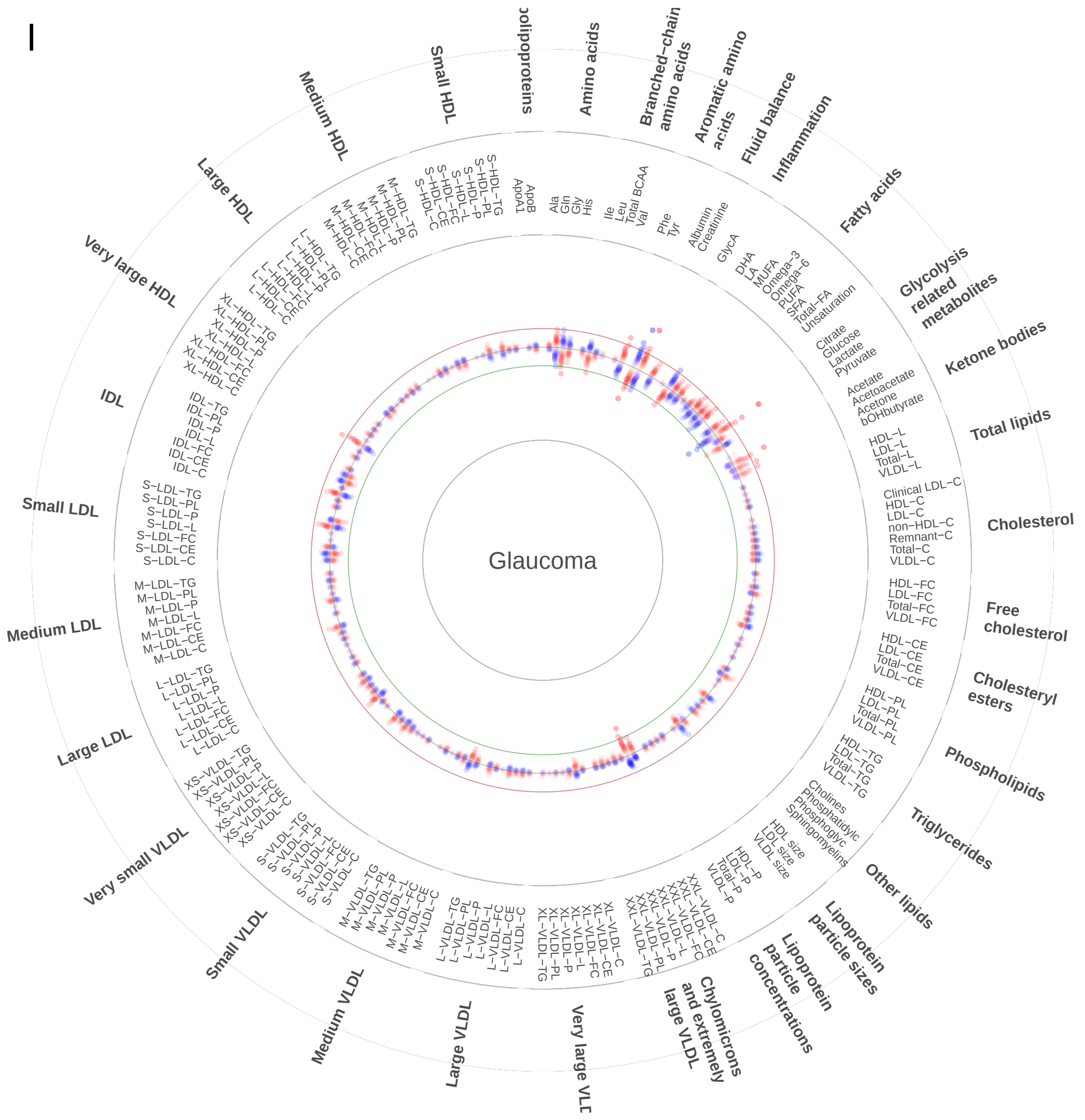


j

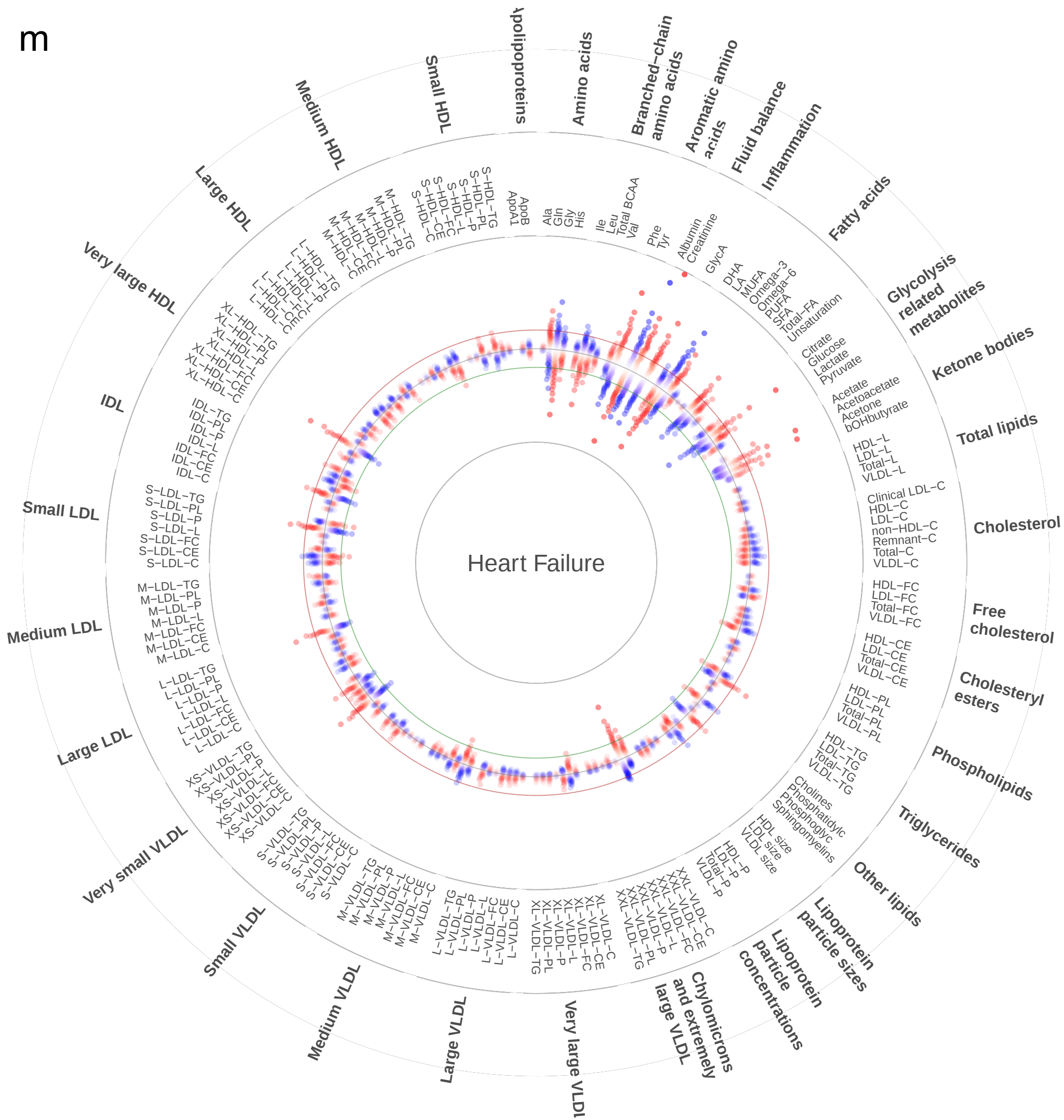


k

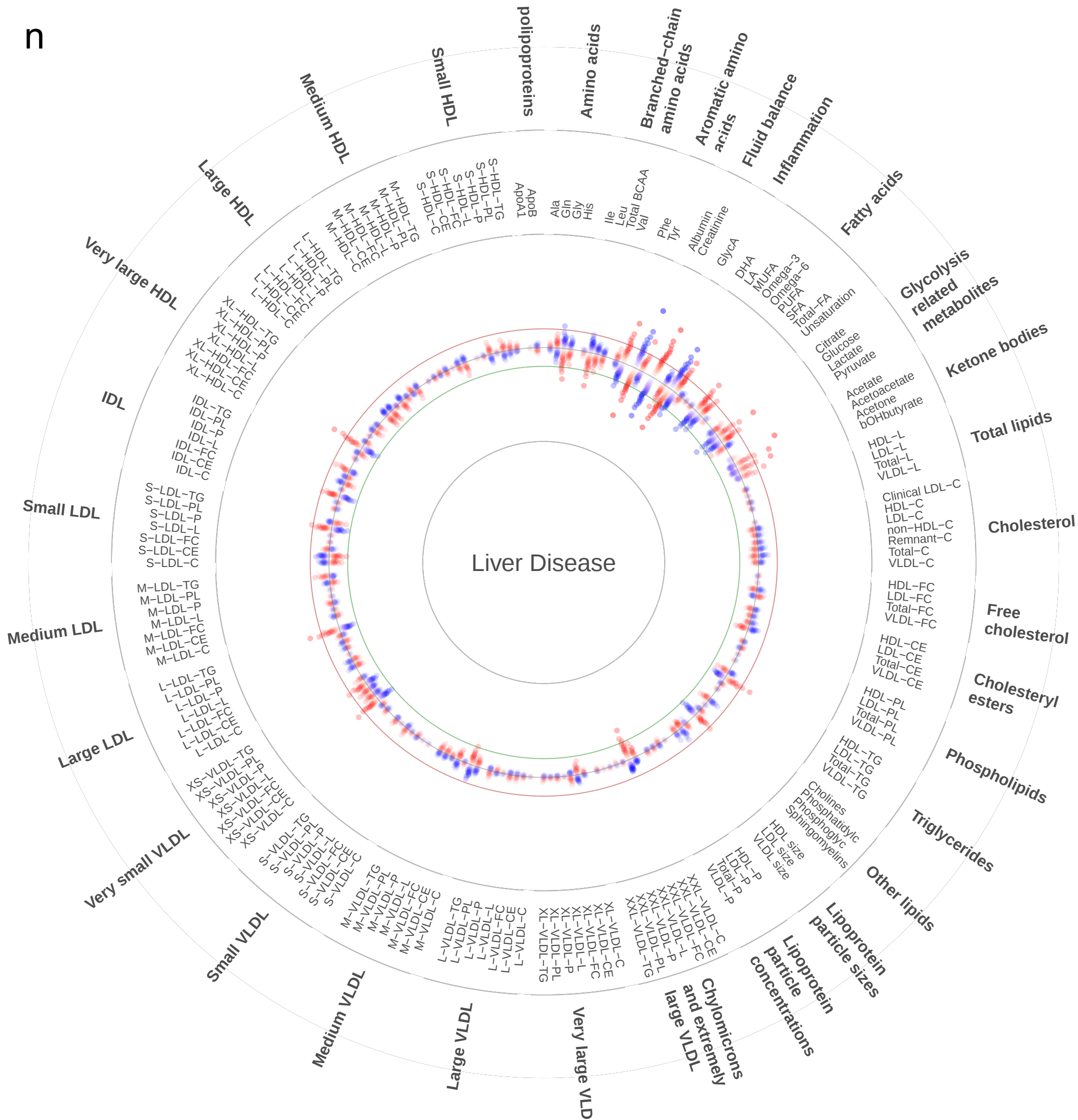




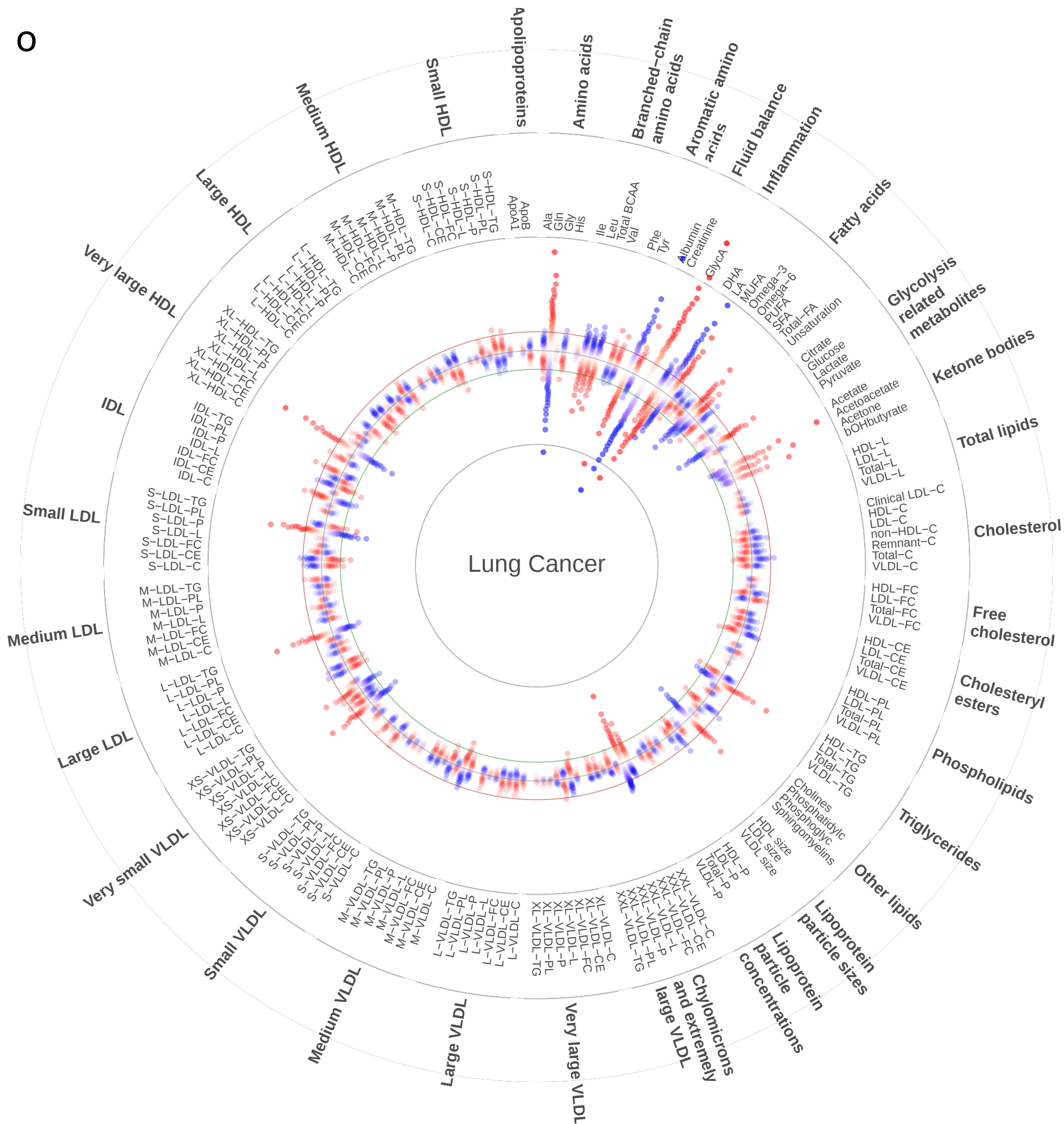
m



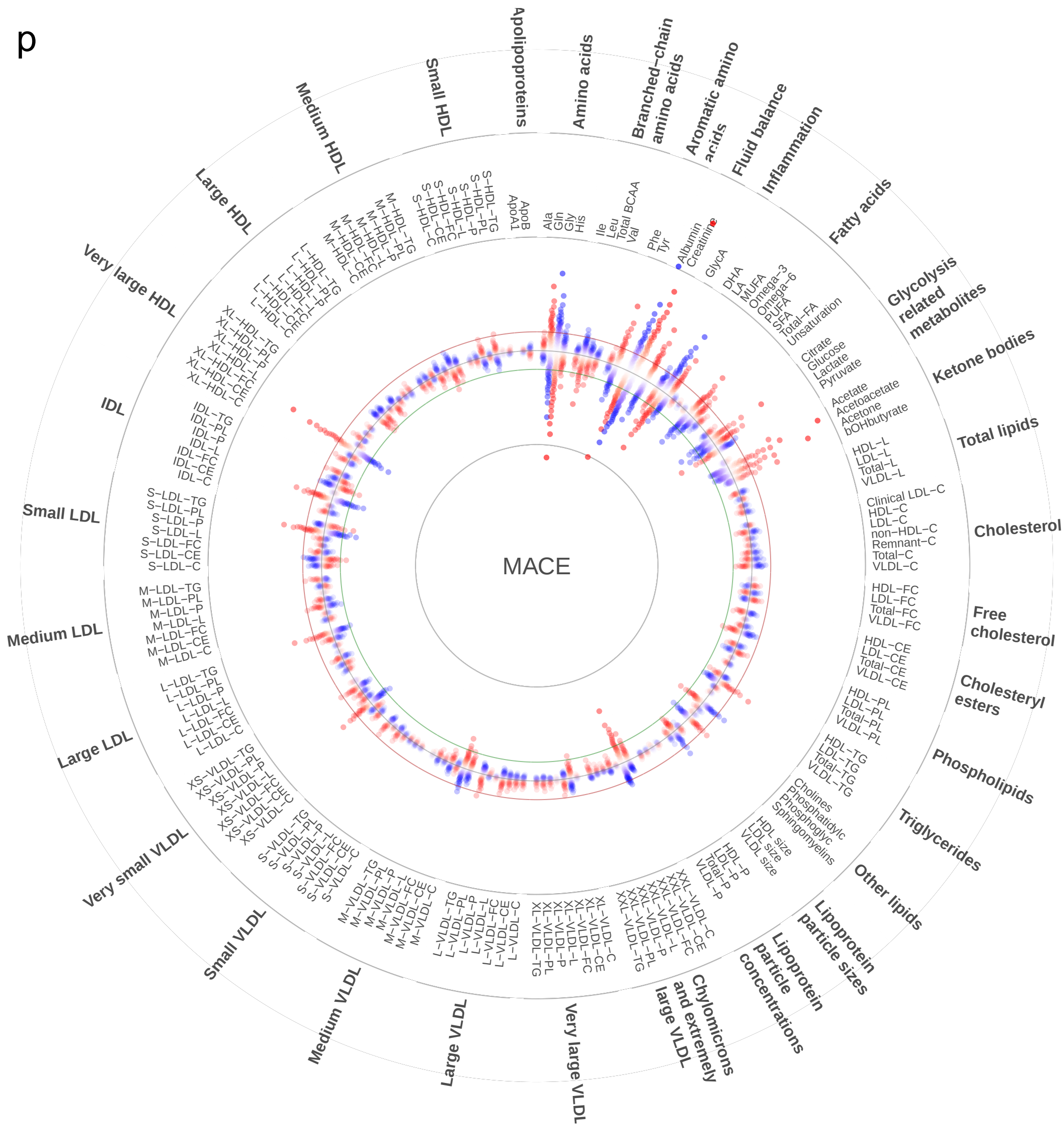
n



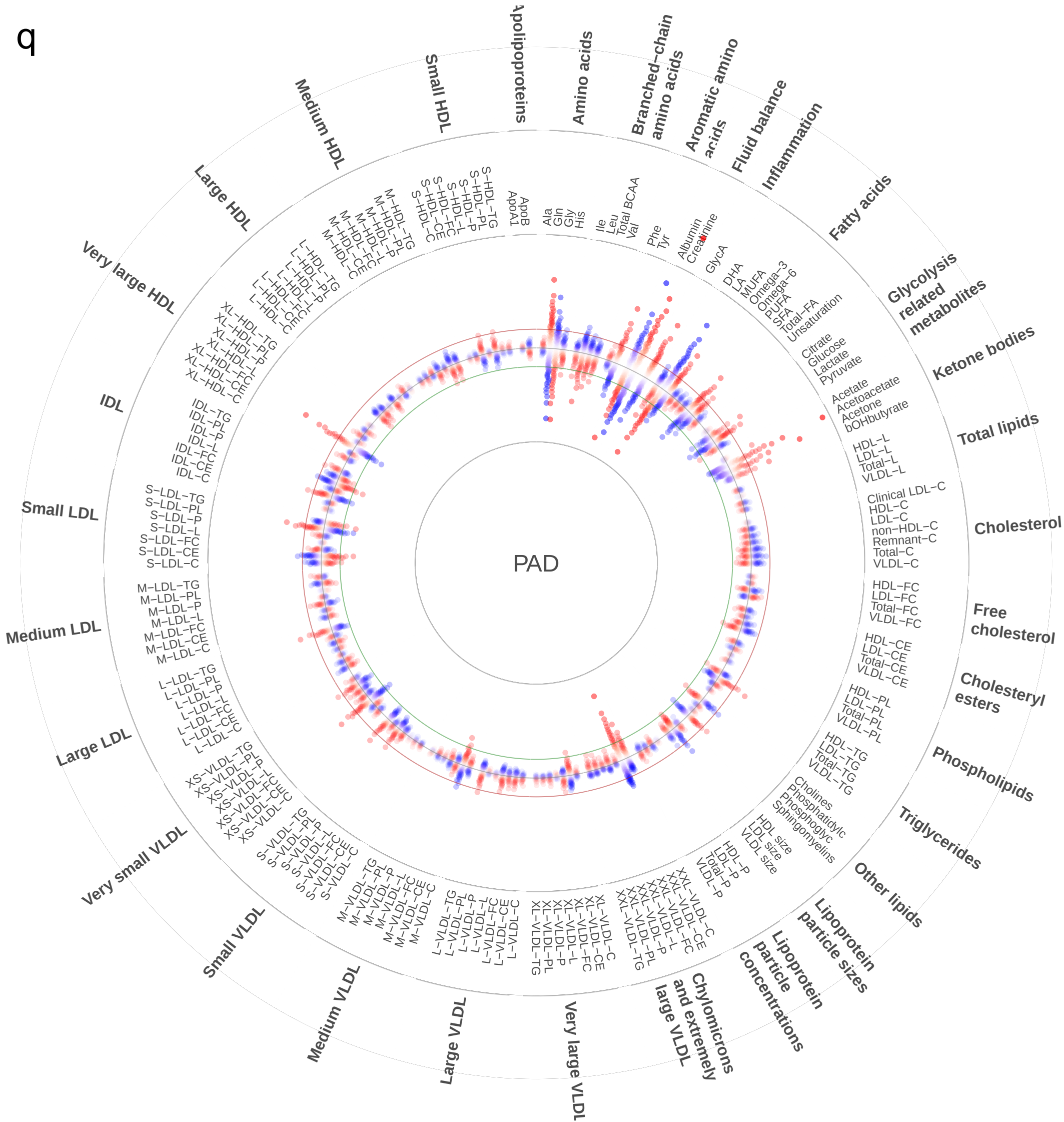
O



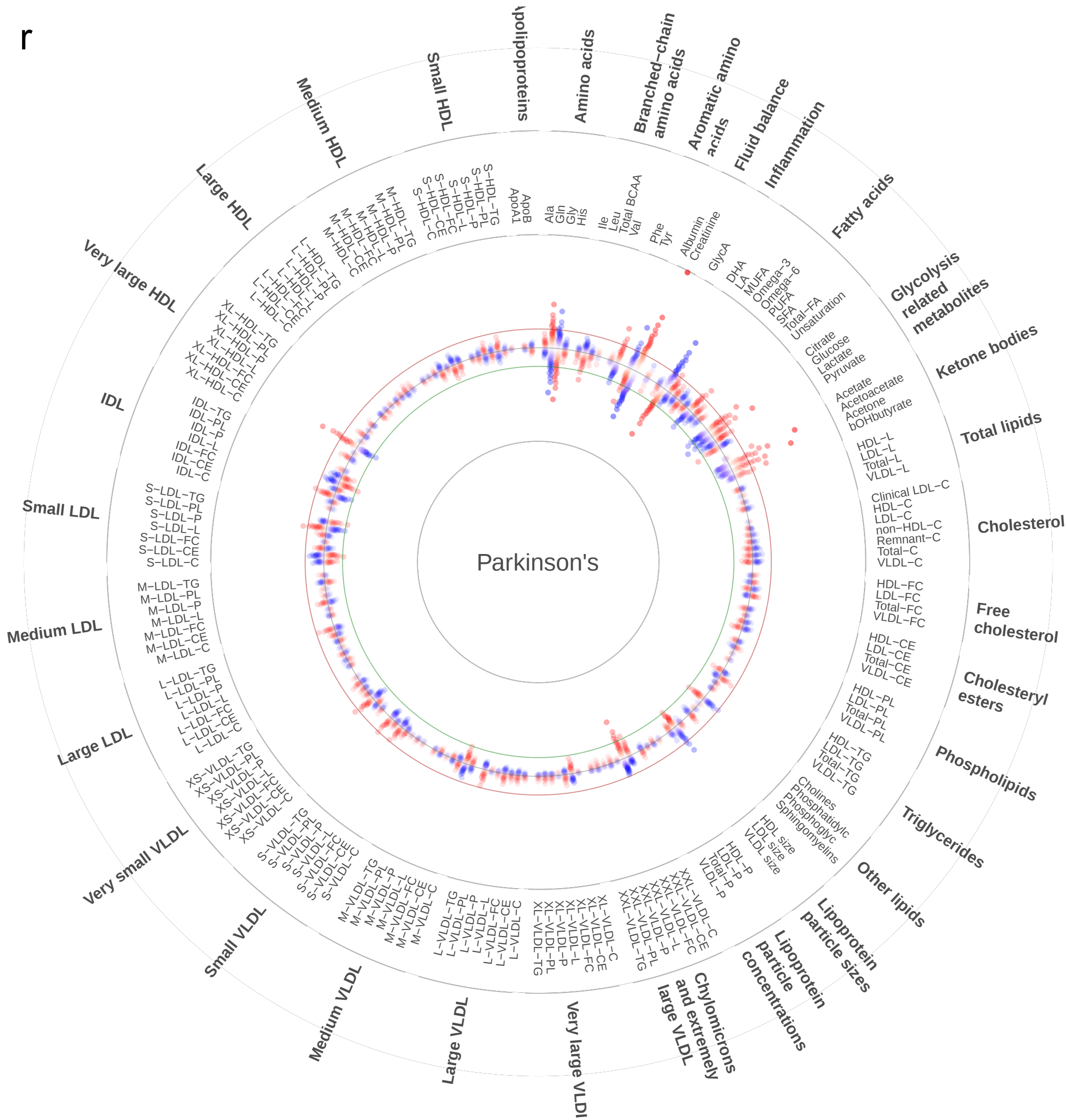
p

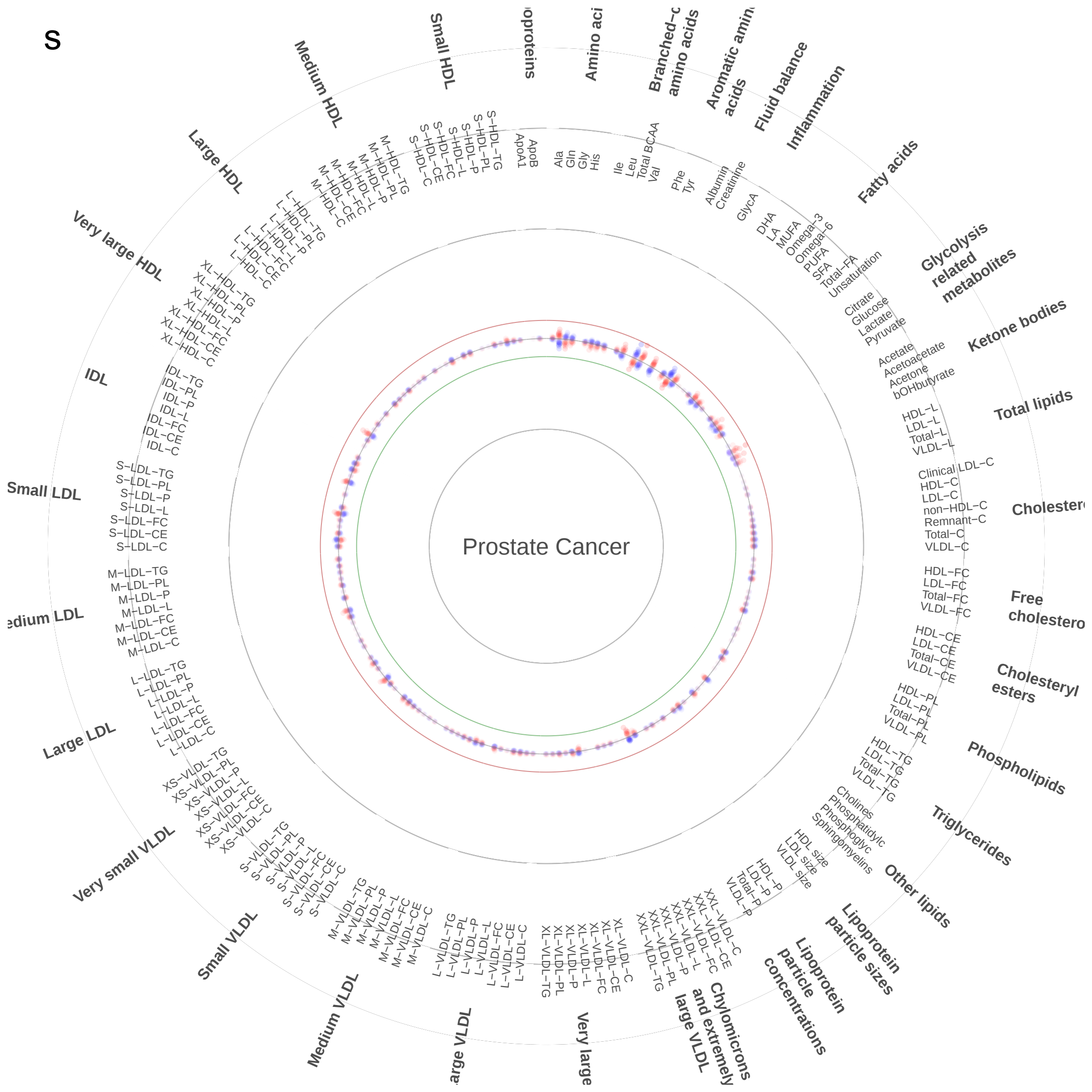


q

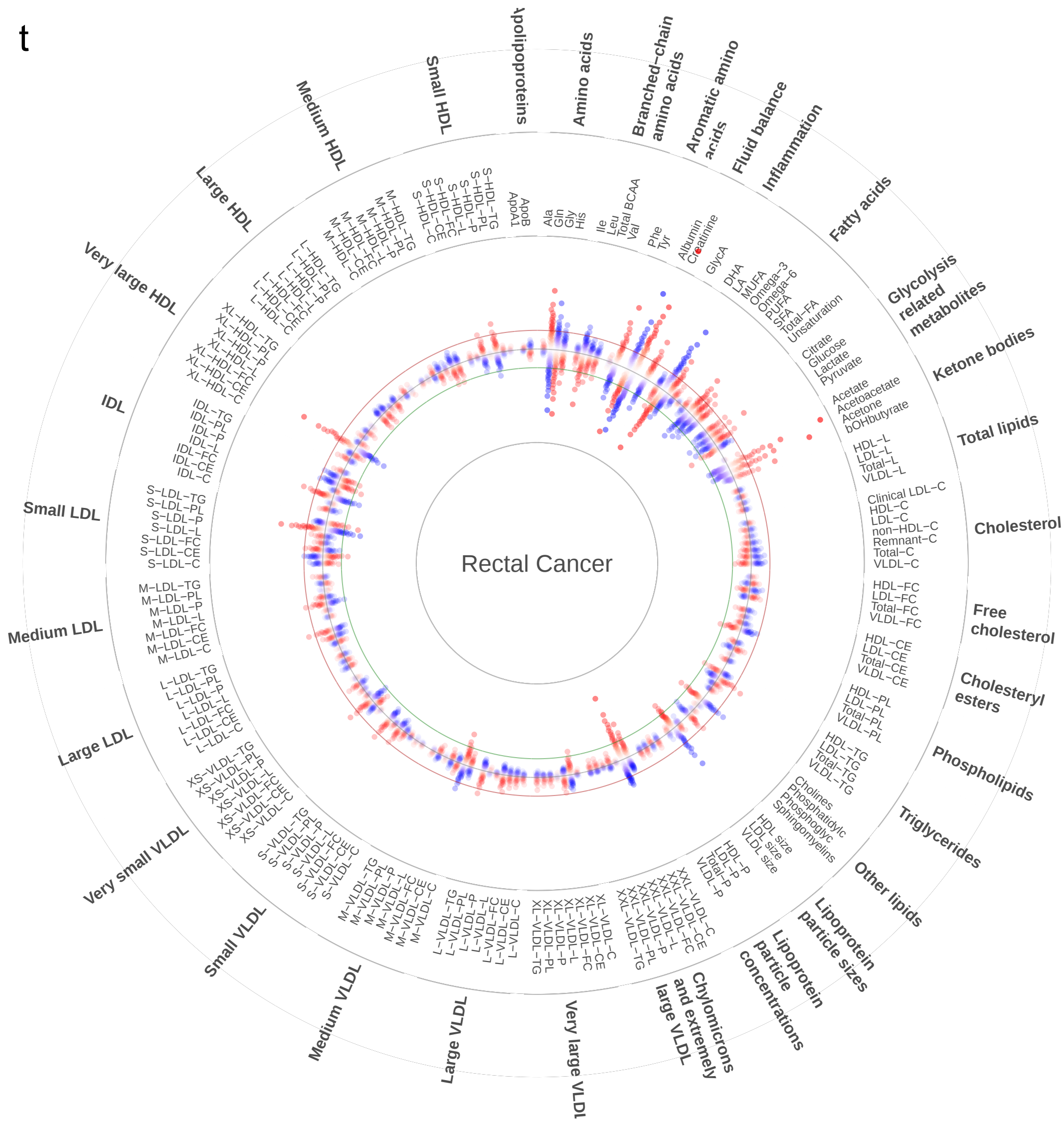


r

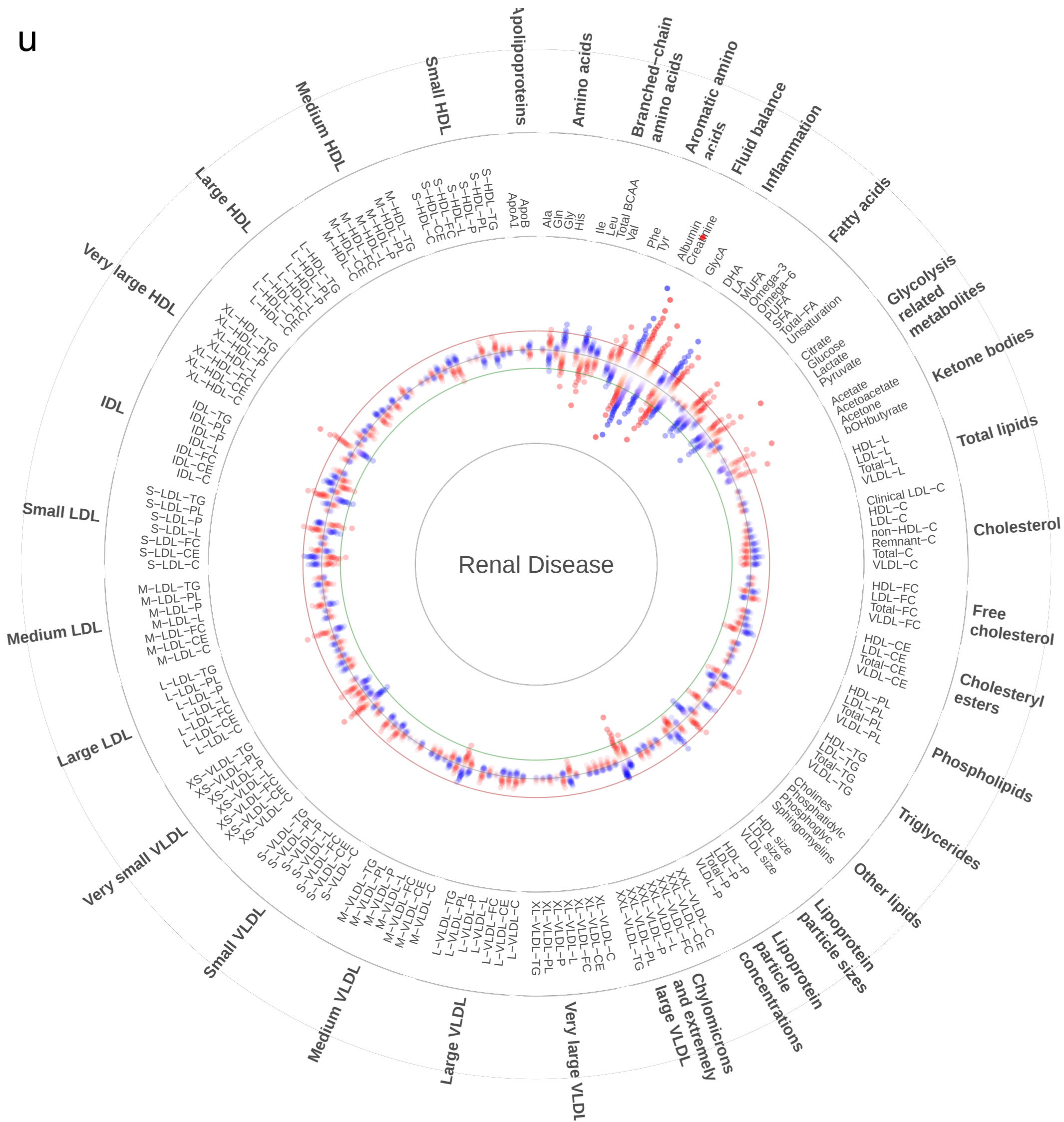




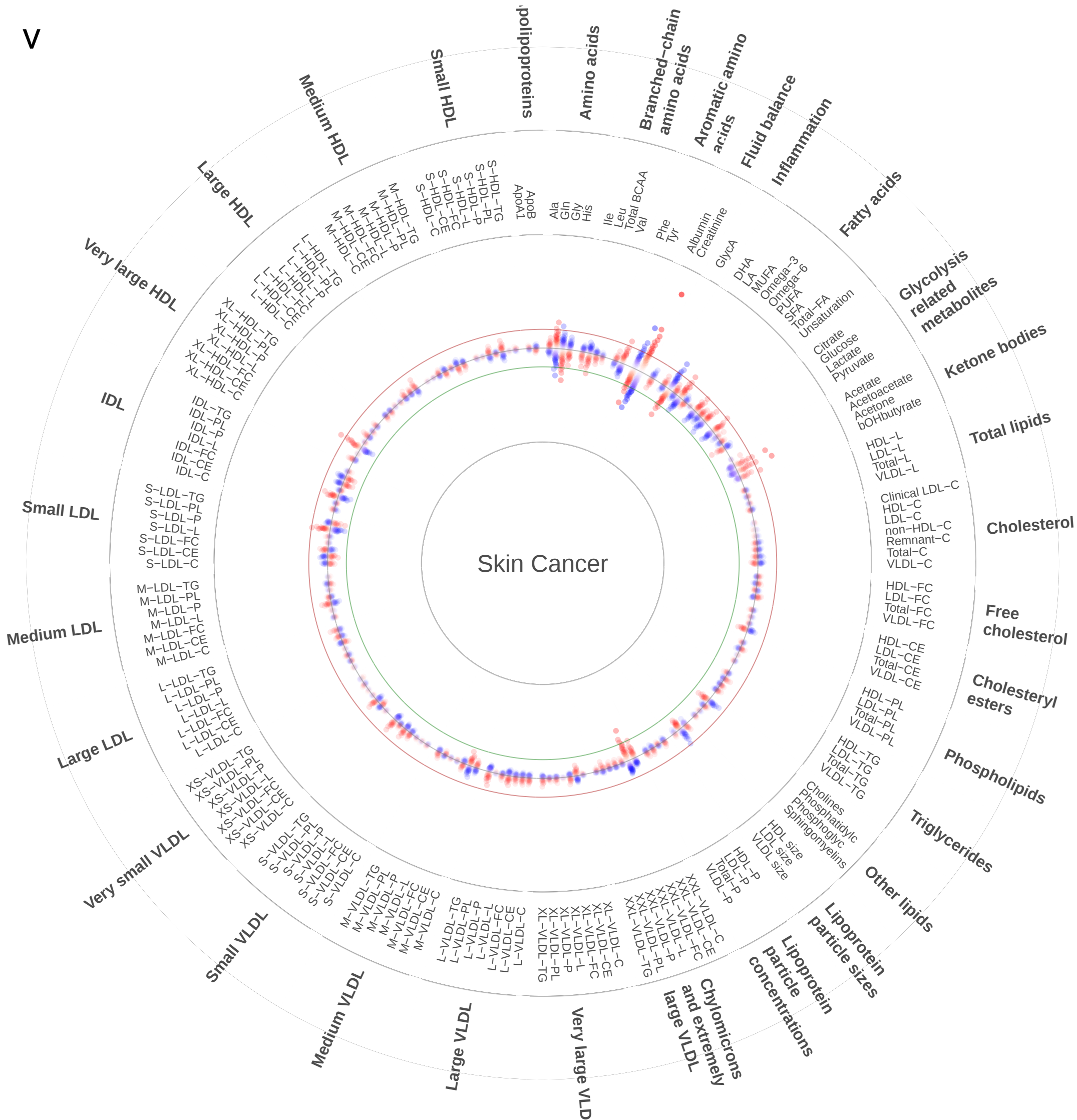
t



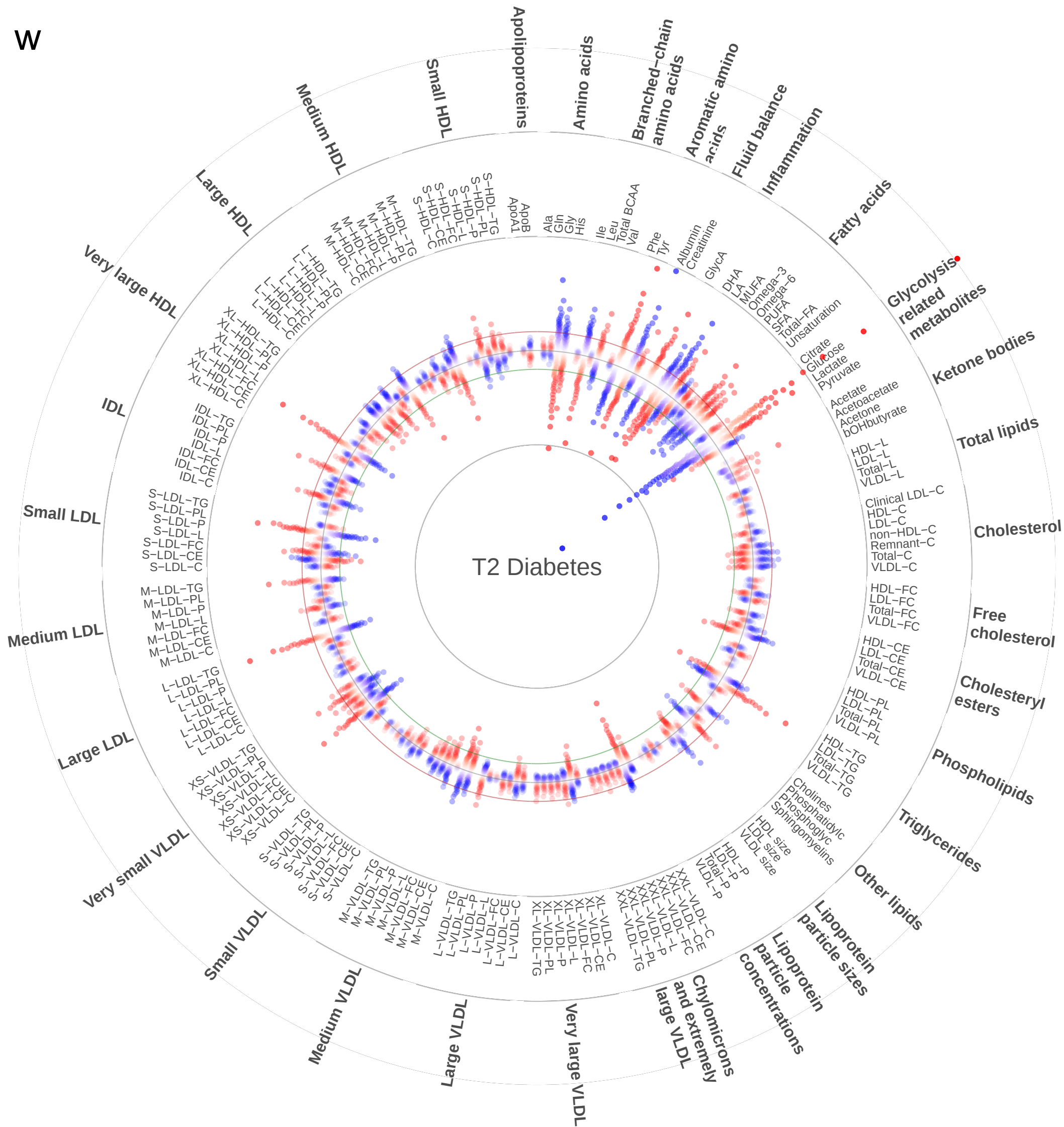
u



V



W



X

